

The MMO Problem

Oscar García-Morchón
 Ronald Rietman
 Ludo Tolhuizen
 Philips Research
 Eindhoven, The Netherlands

Domingo Gómez
 Jaime Gutiérrez
 Universidad de Cantabria
 Santander, Spain

Abstract

We consider a two polynomials analogue of the polynomial interpolation problem. Namely, we consider the Mixing Modular Operations (MMO) problem of recovering two polynomials $f \in \mathbb{Z}_p[x]$ and $g \in \mathbb{Z}_q[x]$ of known degree, where p and q are two (un)known positive integers, from the values of $f(t) \bmod p + g(t) \bmod q$ at polynomially many points $t \in \mathbb{Z}$. We show that if p and q are known, the MMO problem is equivalent to computing a close vector in a lattice with respect to the infinity norm. We also implemented in the SAGE system a heuristic polynomial-time algorithm. If p and q are kept secret, we do not know how to solve this problem. This problem is motivated by several potential cryptographic applications.

1. Introduction

For integer x and integer $p \geq 2$, we denote by $\langle x \rangle_p$ the remainder of dividing x by p . Stated differently,

$$0 \leq \langle x \rangle_p \leq p - 1 \text{ and } x \equiv \langle x \rangle_p \bmod p.$$

The set $\{0, 1, \dots, p-1\}$ can be identified with \mathbb{Z}_p , the ring of integers modulo p . Conversely, \mathbb{Z}_p can be considered as a subset of \mathbb{Z} . This allows us to interpret functions on \mathbb{Z}_p as polynomials evaluated modulo p on the set $\{0, 1, \dots, p-1\}$ and to extend the domain of these polynomials to \mathbb{Z} . Furthermore it allows us to add polynomials over several different rings $\mathbb{Z}_p, \mathbb{Z}_q, \dots$ for different values of the moduli p, q, \dots . This addition we denote by the term Mixing of Modular Operations.

Here we study a variant of the very well known polynomial interpolation problem, where the function to be interpolated is the sum of two polynomials reduced modulo two different unknown numbers p and q .

Problem 1. *Let $p \neq q$ be two positive unknown integers and c another positive integer. Let the function $h : \mathbb{Z} \rightarrow \mathbb{Z}$ be the sum of two unknown reduced polynomials $h(x) = \langle f(x) \rangle_p + \langle g(x) \rangle_q$ for some polynomials $f \in \mathbb{Z}_p[x]$, $g \in \mathbb{Z}_q[x]$ of degree at most α , where α is known. Suppose that the set*

$$J = \{(x_1, h(x_1)), \dots, (x_c, h(x_c))\}$$

*is known, where $x_i \in \mathbb{Z}$, for $i = 1, \dots, c$. The **MMO problem** is to recover p and q and the polynomials f and g .*

This problem seems to be difficult to solve even for very small polynomial degrees and, in fact, we and other colleagues have not managed. For the single-polynomial analogue of Problem 1, we refer to related work in [1]. The main motivation to study this computational problem arises in potential applications to cryptography [2]. Since we could not obtain a solution for the above problem, this paper mainly devotes its attention to a simplified problem statement in which p and q are known.

Problem 2. Let $p \neq q$ be two known positive integers and c another positive integer. Let the function $h : \mathbb{Z} \rightarrow \mathbb{Z}$ be the sum of two unknown reduced polynomials $h(x) = \langle f(x) \rangle_p + \langle g(x) \rangle_q$ for some polynomials $f \in \mathbb{Z}_p[x]$, $g \in \mathbb{Z}_q[x]$ of degree at most α . Suppose that the set

$$J = \{(x_1, h(x_1)), \dots, (x_c, h(x_c))\}$$

is known, where $x_i \in \mathbb{Z}$, for $i = 1, \dots, c$. The **MMO problem with known moduli** is to recover the polynomials f and g .

This is a natural extension of the well known polynomial interpolation problem. Our results show that if c is big enough compared to α , and the points x_1, \dots, x_c are randomly drawn from a large enough interval, the MMO problem has a unique solution f, g , up to an additive constant.

The paper is organized as follows: Section 2 gives the equivalence of the MMO problem to finding all points in a lattice of dimension $c + 2\alpha$ that are close to a target vector with respect to the infinity norm. Section 3 shows the performance of a Sage implementation of the provided heuristic algorithm. In Section 4, we consider the MMO problem for the case that all arguments x_i lie in a short interval. Section 5 concludes this paper.

2. A general approach

2.1. Preliminaries

This section is devoted to the preliminaries needed to understand the results in the paper. Our purpose is not to give a deep treatment of lattices because these are used in this article only as technical tools. For a nice overview from a cryptographic perspective, we recommend the reader [3]. If the reader interests are nearer to the area of number theory, we recommend [4].

Let $\{\mathbf{a}_1, \dots, \mathbf{a}_d\}$ be a set of linearly independent row vectors in \mathbb{R}^s . The set

$$\mathcal{L} = \{\mathbf{z} : \mathbf{z} = c_1 \mathbf{a}_1 + \dots + c_d \mathbf{a}_d, \quad c_1, \dots, c_d \in \mathbb{Z}\}$$

is called an d -dimensional lattice with basis $\{\mathbf{a}_1, \dots, \mathbf{a}_d\}$.

To each lattice \mathcal{L} one can naturally associate its *volume*

$$\text{Vol}(\mathcal{L}) = (\det(BB^t))^{1/2},$$

where $B \in \mathbb{R}^{d \times s}$ is the matrix with rows $\mathbf{a}_1, \dots, \mathbf{a}_d$. The lattice volume is invariant under unimodular transformations of the basis $\{\mathbf{a}_1, \dots, \mathbf{a}_d\}$.

For a vector \mathbf{u} , let $\|\mathbf{u}\|_\infty$ denote its *infinity norm* and by $\|\mathbf{u}\|_2$ its *Euclidean norm*. It is well known that:

$$\|\mathbf{u}\|_\infty \leq \|\mathbf{u}\|_2 \leq \sqrt{s} \|\mathbf{u}\|_\infty.$$

Any basis of a lattice satisfies

$$\text{Vol}(\mathcal{L}) \leq \prod_{i=1}^d \|\mathbf{a}_i\|_2.$$

The famous Minkowski theorem (see [5, Theorem 5.3.6, page 141]) gives an upper bound on $s_\infty(\mathcal{L})$, the length in infinity-norm of the shortest nonzero vector in any d -dimensional lattice \mathcal{L} , in terms of its volume:

$$s_\infty(\mathcal{L}) = \min \{\|\mathbf{z}\|_\infty : \mathbf{z} \in \mathcal{L} \setminus \{\mathbf{0}\}\} \leq \text{Vol}(\mathcal{L})^{1/d} \quad (1)$$

Denote the number of points of a d -dimensional lattice in \mathbb{R}^d that lie in a measurable subset S of \mathbb{R}^d by $N_{\mathcal{L}}(S)$. Let $C(\mathcal{L})$ be a fundamental cell of \mathcal{L} , with volume $\text{Vol}(\mathcal{L})$. The mean number of lattice points in the shifted set $\mathbf{x} + S$, where $\mathbf{x} \in C(\mathcal{L})$, is given by

$$\frac{1}{\text{Vol}(\mathcal{L})} \int_{C(\mathcal{L})} N_{\mathcal{L}}(\mathbf{x} + S) d^d x = \frac{\text{Vol}(S)}{\text{Vol}(\mathcal{L})}.$$

A similar result appears in [6, Lemma 2, page 27], where the number of lattice points inside a d -dimensional ball of radius r is approximated by the volume of the ball divided by the volume of the lattice.

As in [6, Definition 8, page 27], the *Gaussian heuristic* is to neglect the averaging, and estimate the number of lattice points in S as

$$N_{\mathcal{L}}(S) \approx \frac{\text{Vol}(S)}{\text{Vol}(\mathcal{L})}.$$

Take S to be a d -dimensional hypercube of length $2L$, parallel to the coordinate axes and centered around a lattice point. For S to contain one lattice point, L must be less than $s_{\infty}(\mathcal{L})$. The Gaussian heuristic thus suggests that $(2s_{\infty}(\mathcal{L}))^d > \text{Vol}(\mathcal{L})$, giving a lower bound

$$s_{\infty}(\mathcal{L}) > \frac{1}{2} (\text{Vol}(\mathcal{L}))^{1/d},$$

which is precisely half as big as the rigorous upper bound given by the Minkowski theorem.

Finding the shortest vector in the lattice is a difficult task. Indeed, finding the shortest vector of a lattice for the infinity norm is NP -hard. Fortunately, after the breakthrough in [7], it is possible to find “short” vectors in a lattice, thanks to the concept of *LLL-reduced basis*. For the *LLL*-reduced basis $\mathbf{a}_1, \dots, \mathbf{a}_d$ and its Gram-Schmidt orthogonalization $\mathbf{a}_1^*, \dots, \mathbf{a}_d^*$ there exist real numbers μ_{ij} for $1 \leq j \leq i \leq d$ such that

$$\begin{aligned} |\mu_{ij}| &\leq 1/2, \text{ for } 1 \leq j < i \leq d, \\ \|\mathbf{a}_i^* + \mu_{ii-1}\mathbf{a}_{i-1}^*\|_2^2 &\leq \epsilon \|\mathbf{a}_{i-1}^*\|_2^2, \text{ for } i = 1 \dots, d-1. \end{aligned}$$

for some $\epsilon \in (1/4, 1)$.

Finally, we introduce the following notation. For each real x , we denote by $\lfloor x \rfloor$ the value of x rounded downwards to the closest integer, that is,

$$\lfloor x \rfloor = \max\{m \in \mathbb{Z} \mid m \leq x\}.$$

Lemma 1. *For any integer x and any integer $p > 1$, we have:*

- $\langle x \rangle_p = x - \lfloor x/p \rfloor p$
- *There is a unique integer λ such that $|2x - 2p\lambda - (p-1)| < p$. For this integer it holds that $\lambda = \lfloor x/p \rfloor$.*

Similarly, for an integer vector $\mathbf{x} = (x_1, \dots, x_d)$, $\lfloor \mathbf{x}/p \rfloor$ is equal to the unique integer vector $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_d)$ such that for each component it holds that $|2x_k - 2p\lambda_k - (p-1)| < p$. If \mathbf{e}_d is the vector of length d with all components equal to 1, the latter condition is equivalent to $\|2\mathbf{x} - 2p\boldsymbol{\lambda} - (p-1)\mathbf{e}_d\|_{\infty} < p$.

2.2. Lattice reduction

The next proposition shows that from the values of $h(x) = \langle f(x) \rangle_p + \langle g(x) \rangle_q$ in all integers x , the polynomials $f \in \mathbb{Z}_p[x]$ and $g \in \mathbb{Z}_q[x]$ are determined uniquely up to constant.

Proposition 1. *Let p and q be two positive integers that are relatively prime. Let f, g, u, v be functions from \mathbb{Z} to \mathbb{Z} such that for each integer x ,*

$$\langle f(x) \rangle_p + \langle g(x) \rangle_q = \langle u(x) \rangle_p + \langle v(x) \rangle_q.$$

There exists an integer C such that for each integer x , we have that

$$\langle u(x) \rangle_p = \langle f(x) \rangle_p + C \text{ and } \langle v(x) \rangle_q = \langle g(x) \rangle_q - C.$$

Proof: For each integer x , we have that

$$\langle f(x) \rangle_p - \langle u(x) \rangle_p = \langle v(x) \rangle_q - \langle g(x) \rangle_q.$$

The function $\langle f(x) \rangle_p - \langle u(x) \rangle_p$, which clearly is periodic with period p , thus also is periodic with period q , and thus is periodic with period $\gcd(p, q)=1$, that is, the function is constant. \square

Since it must hold, for every x , that $0 \leq \langle f(x) \rangle_p, \langle u(x) \rangle_p \leq p-1$ and $0 \leq \langle g(x) \rangle_q, \langle v(x) \rangle_q \leq q-1$, it follows that, for all x

$$\max(-\langle f(x) \rangle_p, \langle g(x) \rangle_q - q + 1) \leq C \leq \min(p - 1 - \langle f(x) \rangle_p, \langle g(x) \rangle_q).$$

In particular, C must be equal to 0, and thus the decomposition of the function h must be unique, if there is an x for which $\langle f(x) \rangle_p = \langle g(x) \rangle_q = 0$. That is the reason why we will suppose that $f(0) = g(0) = 0$. Additionally, we are going to suppose $\gcd(p, q) = 1$. Under this condition there exist integers μ_1 and μ_2 such that $\mu_1 p + \mu_2 q = 1$. We want to mention now that if p is much larger than q , then the MMO problem with known moduli can be easily transformed in a *noisy polynomial interpolation problem* (see [8]), where the evaluation of the polynomial g modulo q can be seen as random “noise” and the attacker tries to recover f . Rigorous bounds for the noise of the results in [8] depend heavily on the performance of finding a close vector in the lattice and it seems that there is some gap between the theoretic results and the practical experiments. For this paper, we focus on the case that p and q have approximately the same number of bits.

Without loss of generality, the expression of the polynomials f, g is

$$f(x) = \sum_{k=1}^{\alpha} r_k x^k, \quad g(x) = \sum_{k=1}^{\alpha} t_k x^k,$$

where $r_k, t_k \in \mathbb{Z}$ and $|r_k| < p/2$, $|t_k| < q/2$ for $k = 1, \dots, \alpha$.

We will show that the MMO problem is related to finding a short vector in a lattice. For that, we need the following definitions:

From x_1, \dots, x_c we build the Vandermonde matrix \mathbf{V} of size $\alpha \times c$ as

$$\mathbf{V} = \begin{pmatrix} x_1 & x_2 & \cdots & x_c \\ x_1^2 & x_2^2 & \cdots & x_c^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^\alpha & x_2^\alpha & \cdots & x_c^\alpha \end{pmatrix}.$$

Also, for integer x we denote by $h(x) = \langle f(x) \rangle_p + \langle g(x) \rangle_q$. The MMO problem can now be formulated as follows: given the vector \mathbf{h} of which the components are the function values $\mathbf{h} = (h(x_1), \dots, h(x_c))$, find integer vectors \mathbf{r}, \mathbf{t} of length α such that $\|\mathbf{r}\|_\infty < p/2$, $\|\mathbf{t}\|_\infty < q/2$ and

$$\mathbf{h} = \langle \mathbf{rV} \rangle_p + \langle \mathbf{tV} \rangle_q = \mathbf{rV} - p[\mathbf{rV}/p] + \mathbf{tV} - q[\mathbf{tV}/q]$$

where all the modulo and rounding operations act component-wise.

Using Lemma 1, it is clear that the MMO problem can be restated as follows: given \mathbf{h} , find integer row vectors \mathbf{r}, \mathbf{t} of length α and λ_1, λ_2 of length c such that

$$\mathbf{h} = \mathbf{rV} - p\lambda_1 + \mathbf{tV} - q\lambda_2, \quad (2)$$

and

$$\left\| \frac{\mathbf{r}\mathbf{V}}{p} - \boldsymbol{\lambda}_1 - \frac{(p-1)\mathbf{e}_c}{2p} \right\|_\infty < \frac{1}{2}, \quad \left\| \frac{\mathbf{t}\mathbf{V}}{q} - \boldsymbol{\lambda}_2 - \frac{(q-1)\mathbf{e}_c}{2q} \right\|_\infty < \frac{1}{2}. \quad (3)$$

The inequalities in (3) embody the constraints that the vectors $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2$ are the result of the rounding operation.

We concatenate the vectors $\mathbf{r}, \mathbf{t}, \boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ vector \mathbf{x} of length $2(c + \alpha)$:

$$\mathbf{x} = (\mathbf{r}, \mathbf{t}, -\boldsymbol{\lambda}_1, -\boldsymbol{\lambda}_2)$$

and define a matrix \mathbf{A} of size $2(c + \alpha) \times c$ as a vertical concatenation of 2 copies of \mathbf{V} and 2 instances of the $c \times c$ identity matrix \mathbf{I}_c multiplied by p, q respectively:

$$\mathbf{A} = \begin{pmatrix} \mathbf{V} \\ \mathbf{V} \\ p\mathbf{I}_c \\ q\mathbf{I}_c \end{pmatrix},$$

so that equation (2) becomes

$$\mathbf{h} = \mathbf{x}\mathbf{A} \quad (4)$$

Furthermore we define the matrix \mathbf{B} of size $2(c + \alpha) \times 2(c + \alpha)$ as the block matrix

$$\mathbf{B} = \begin{pmatrix} \mathbf{I}_\alpha/p & \mathbf{0}_{\alpha \times \alpha} & \mathbf{V}/p & \mathbf{0}_{\alpha \times \alpha} \\ \mathbf{0}_{\alpha \times \alpha} & \mathbf{I}_\alpha/q & \mathbf{0}_{\alpha \times c} & \mathbf{V}/q \\ \mathbf{0}_{c \times \alpha} & \mathbf{0}_{c \times \alpha} & \mathbf{I}_c & \mathbf{0}_{c \times c} \\ \mathbf{0}_{c \times \alpha} & \mathbf{0}_{c \times \alpha} & \mathbf{0}_{c \times c} & \mathbf{I}_c \end{pmatrix}$$

and the vector \mathbf{u} of length $2(c + \alpha)$ as

$$\mathbf{u} = (\underbrace{0, \dots, 0}_{2\alpha}, \frac{p-1}{2p}\mathbf{e}_c, \frac{q-1}{2q}\mathbf{e}_c).$$

Now the inequalities (3) and $\|\mathbf{r}\|_\infty < p/2, \|\mathbf{t}\|_\infty < q/2$ are equivalent to the single inequality

$$\|\mathbf{x}\mathbf{B} - \mathbf{u}\|_\infty < \frac{1}{2}. \quad (5)$$

So finding a solution to the MMO problem is equivalent to finding an integer solution of equation (4) that satisfies the constraint from inequality (5).

Let \mathbf{x}_0 be an arbitrary integer solution to equation (4), for example we can take $\mathbf{x}_0 = (\underbrace{0, \dots, 0}_{2\alpha}, \mu_1\mathbf{h}, \mu_2\mathbf{h})$. Every integer solution \mathbf{x} of equation (4) can now be written as $\mathbf{x} = \mathbf{x}_0 + \mathbf{y}$, where $\mathbf{y}\mathbf{A} = \mathbf{0}$. Thus \mathbf{y} lies in the left integer kernel of \mathbf{A} , which is spanned by the rows of the matrix

$$\mathbf{K} = \begin{pmatrix} \mathbf{I}_\alpha & -\mathbf{I}_\alpha & \mathbf{0}_{\alpha \times c} & \mathbf{0}_{\alpha \times c} \\ \mathbf{0}_{\alpha \times \alpha} & \mathbf{I}_\alpha & -\mu_1\mathbf{V} & -\mu_2\mathbf{V} \\ \mathbf{0}_{c \times \alpha} & \mathbf{0}_{c \times \alpha} & q\mathbf{I}_c & -p\mathbf{I}_c \end{pmatrix},$$

so $\mathbf{y} = \mathbf{w}\mathbf{K}$ with $\mathbf{w} \in \mathbb{Z}^{2\alpha+c}$. Substituting this into equation (5), we obtain

$$\|\mathbf{w}\mathbf{KB} - (\mathbf{u} - \mathbf{x}_0\mathbf{B})\|_\infty < \frac{1}{2}.$$

In other words, we are looking for vectors in the lattice \mathcal{L} spanned by the rows of the matrix

$$\mathbf{C} = \mathbf{KB} = \begin{pmatrix} \mathbf{I}_\alpha/p & -\mathbf{I}_\alpha/q & \mathbf{V}/p & -\mathbf{V}/q \\ \mathbf{0}_{\alpha \times \alpha} & \mathbf{I}_\alpha/q & -\mu_1\mathbf{V} & \mu_1p\mathbf{V}/q \\ \mathbf{0}_{c \times \alpha} & \mathbf{0}_{c \times \alpha} & q\mathbf{I}_c & -p\mathbf{I}_c \end{pmatrix}$$

that have distance less than $1/2$ in infinity norm to the vector $\mathbf{u} - \mathbf{x}_0\mathbf{B}$.

The main idea of the lattice reduction technique is to show that the close vector is unique. Suppose we have two lattice vectors \mathbf{z}_1 and $\mathbf{z}_2 \in \mathcal{L}$ satisfying

$$\|\mathbf{z}_i - (\mathbf{u} - \mathbf{x}_0\mathbf{B})\|_\infty < \frac{1}{2}, \quad i = 1, 2,$$

then $\mathbf{z} = \mathbf{z}_1 - \mathbf{z}_2 \in \mathcal{L}$ and $\|\mathbf{z}\|_\infty < 1$.

Note that the fourth block column of \mathbf{C} is equal to $-p/q$ times the third. This implies that for each $\mathbf{z} \in \mathcal{L}$, we have that $\|\mathbf{z}\|_\infty = \|\mathbf{z}'\|_\infty$, where $\mathbf{z}' \in \mathbb{Q}^{2\alpha+c}$ is obtained from \mathbf{z} by deleting the last block of c coordinates if $p < q$ and the third block if $q < p$. Deleting the corresponding block column from \mathbf{C} gives a square matrix \mathbf{C}' ; the $(2\alpha + c)$ -dimensional lattice of which the rows of \mathbf{C}' are a basis is denoted \mathcal{L}' . Then

$$\text{Vol}(\mathcal{L}') = |\det(\mathbf{C}')| = \frac{\max(p, q)^c}{(pq)^\alpha}.$$

The Gaussian heuristic suggests that a d -dimensional lattice \mathcal{L}' with volume $\text{Vol}(\mathcal{L}')$ is unlikely to have a nonzero vector which is substantially shorter (in infinity norm) than $(1/2)\text{Vol}(\mathcal{L}')^{1/d}$. Thus, if $\text{Vol}(\mathcal{L}') > 2^{2\alpha+c}$ it is likely that the close vector is unique. When p and q have similar magnitude, we therefore conclude that if c is somewhat larger than 2α , it is likely that the MMO problem can be solved.

Conversely, with elementary methods we can show that if p and q have similar magnitude, then reconstruction of (f, g) requires that on average, c is at least 2α . Indeed, the number of pairs of polynomials (f, g) equals $(pq)^\alpha$; the number of sequences of function values in c integers equals $(p + q - 1)^c$. Hence, if $(pq)^\alpha > (p + q - 1)^c$, then there exists a sequence of function values that can be generated by more than one pair (f, g) of polynomials. The following proposition gives a slightly stronger result.

Proposition 2. *If p and q have similar magnitude, then on average the minimum number of required values to compute the polynomials $\langle f(X) \rangle_p$ and $\langle g(X) \rangle_q$ is at least 2α .*

Proof: Let x_1, \dots, x_c be integers. For $\mathbf{y} \in Y = \{0, 1, \dots, p + q - 2\}^c$, we define

$$N(\mathbf{y}) = |\{(f, g) \in \mathbb{Z}_p[x] \times \mathbb{Z}_q[x] \mid \deg(f) \leq \alpha, \deg(g) \leq \alpha, f(0) = g(0) = 0 \text{ and} \\ \langle f(x_i) \rangle_p + \langle g(x_i) \rangle_q = y_i \text{ for } 1 \leq i \leq c\}|.$$

Of course, we have that

$$\sum_{\mathbf{y} \in Y} N(\mathbf{y}) = p^\alpha q^\alpha.$$

We assume the polynomials f and g are chosen uniformly and independently. Then the probability $p(\mathbf{y})$ to observe $\mathbf{y} \in Y$ equals $N(\mathbf{y})/(pq)^\alpha$. The expected number \mathcal{E} of pairs of polynomials (f, g) matching $\mathbf{y} \in Y$ thus satisfies

$$\mathcal{E} = \sum_{\mathbf{y} \in Y} N(\mathbf{y})p(\mathbf{y}) = \frac{1}{(pq)^\alpha} \sum_{\mathbf{y} \in Y} N(\mathbf{y})^2 \geq \frac{1}{(pq)^\alpha} \frac{\left(\sum_{\mathbf{y} \in Y} N(\mathbf{y})\right)^2}{|Y|} = \frac{(pq)^\alpha}{(p + q - 1)^c},$$

where the inequality sign follows from the Cauchy-Schwarz inequality.

Consequently, if $c \leq 2\alpha - 1$, then $\mathcal{E} \geq \frac{(pq)^\alpha}{(p + q - 1)^{2\alpha - 1}} \geq (p + q) \left(\frac{pq}{(p + q)^2}\right)^\alpha$. And so, writing $q = p(1 + \epsilon)$, we have that $\mathcal{E} \geq p(2 + \epsilon) \left(\frac{1 + \epsilon}{(2 + \epsilon)^2}\right)^\alpha$.

For sufficiently small ϵ , we thus have that $\mathcal{E} > 1$. \square

In the next section we provide the details of the resulting algorithm and the performance of our Sage implementation.

3. The algorithm and its implementation

The basic structure of the algorithm is the following:

Algorithm 1 Algorithm to solve MMO problem

Require: Set J and p, q

Ensure: $\langle f(X) \rangle_p$ and $\langle g(X) \rangle_q$.

Generate vectors, $\mathbf{h}, \mathbf{x}_0, \mathbf{u}$ and matrices $\mathbf{B}, \mathbf{K}, \mathbf{C}$ as defined in Section 2.

Use a Closest Vector algorithm to find \mathbf{x}' .

return the polynomials with coefficients equal to the first 2α components of vector \mathbf{x}'

This is the pseudocode of the algorithm we have used to compute a close vector which is called the *Babai Nearest Plane Algorithm*, see [9]:

Algorithm 2 Babai Nearest Plane algorithm

Require: Basis given as a matrix \mathbf{B}, \mathbf{t}

Ensure: A vector $\mathbf{u} \in \mathcal{L}(\mathbf{B})$, such that $\|\mathbf{u} - \mathbf{t}\|_2 \leq 2^{d/2} \min\{\|\mathbf{v} - \mathbf{t}\|_2 \mid \mathbf{v} \in \mathcal{L}(\mathbf{B})\}$

Run LLL algorithm on matrix \mathbf{B} with standard $\epsilon = 3/4$

$\mathbf{b} = \mathbf{t}$

for j from n to 1 **do**

$c_j = \lceil \frac{\mathbf{b}\mathbf{b}_j}{\|\mathbf{b}_j\|_2^2} \rceil$

$\mathbf{b} = \mathbf{b} - c_j \mathbf{b}_j$

end for

return $\mathbf{b} - \mathbf{t}$

We have implemented our algorithm for solving the MMO problem in the Sage system, including the Babai algorithm.

Babai Nearest Plane algorithm finds a close vector with respect to the Euclidean norm. The closest vector with respect to the infinity norm can be found doing the following computations:

- Calculate a *LLL*-reduced basis $\mathbf{a}_1, \dots, \mathbf{a}_d$.
- Calculate a close vector \mathbf{b} using the Babai Nearest Plane algorithm.
- Take the vector \mathbf{b}' that minimizes $\|\mathbf{t} - \mathbf{b}'\|_\infty$ where \mathbf{b}' belongs to the following set,

$$\{\mathbf{b}' \mid \mathbf{b}' = \mathbf{b} + \sum_{i=1}^d C_i \mathbf{a}_i, |C_i| \leq \sqrt{d} 2^{(i-1)/2}, i = 1, \dots, d\}.$$

The fact that this returns the closest vector with respect to the infinity norm comes from [7, Proposition 1.6] and the proof of [7, Proposition 1.11].

To test when the algorithm to solve MMO works, we use an indirect method. We take the lattice defined by the rows of \mathbf{C} and check for the shortest vector. If this vector has norm bigger than 1, then we know that the algorithm will work and in other case, we suppose that it fails.

In this way, we will count as fails many cases where the algorithm could possibly work. However, implementations show that, even in these conditions, the algorithm for solving MMO seems to work in most of the cases. To be more precise, selecting uniformly at random $c = 2\alpha$ values $x_i \in [1, p]$ the algorithm was successful in 100% of the cases with 200-bit number p . This confirms that $c = 2\alpha$ is indeed the natural threshold for the algorithm.

However the performance changes if the values are selected from a small interval $[1, p^{1/K}]$ for big K . If K is smaller than α then it is possible to recover some of the coefficients of

the polynomials. More precisely, the algorithm recovers the coefficients of the polynomials of the monomials of degree greater than K .

This fact is interesting because of the design of the HIMMO key generation system [2] and it is analyzed in detail in next section.

4. Restriction to small arguments

In Proposition 1 we showed that f and g are determined up to a constant if $h(x) = \langle f(x) \rangle_p + \langle g(x) \rangle_q$ for all $x \in \mathbb{Z}$. This constant can be fixed by setting $f(0) = g(0) = 0$. However, in cryptographic applications, values of x that can be used are from a smaller interval: $0 \leq x < w$, where $w \approx (\min(p, q))^{1/K}$ for some $K \geq 1$. If we are interested only in the function h on this short interval, then the reconstruction is typically far from unique. In fact, let $C \in \mathbb{Q}[x]$ be a polynomial of degree at most K that takes integer values for all integer arguments, i.e., C is an integer linear combination of binomial coefficients:

$$C(x) = \sum_{k=0}^K C_k \binom{x}{k}, \quad C_0, \dots, C_K \in \mathbb{Z}.$$

If $\gcd(p, K!) = \gcd(q, K!) = 1$, the factorials $2!, 3!, \dots, K!$ have inverses modulo p and modulo q , so we can define polynomials $c_p \in \mathbb{Z}_p[x]$ and $c_q \in \mathbb{Z}_q[x]$ of degree at most K , such that for all integer x :

$$\langle c_p(x) \rangle_p = \langle C(x) \rangle_p \text{ and } \langle c_q(x) \rangle_q = \langle C(x) \rangle_q.$$

If it holds that C is small on $[0, w)$, in the sense that

$$0 \leq \langle f(x) \rangle_p + C(x) \leq p-1 \text{ and } 0 \leq \langle g(x) \rangle_q - C(x) \leq q-1 \text{ for all integer } x \in [0, w),$$

then $f + c_p$ and $g - c_q$ decompose h on $[0, w)$.

If all short lattice vectors correspond to such pairs $(c_p, -c_q)$, then all lattice points close to our target vector correspond to polynomials (f, \tilde{g}) that also decompose h . In other words: though we cannot reconstruct f and g , we can interpolate h correctly.

Note that our previous analysis based on lattice volumes and the Gaussian heuristic failed to see the short vectors that correspond to the polynomials $C(x)$. This should not be surprising: the lattice volume is independent of the values x_1, \dots, x_c , and these short vectors appear only if $0 \leq x_i < w$ for $i = 1, 2, \dots, c$. The numerical experiments show that the Gaussian heuristic is not valid for \mathcal{L}' when the x_i are from an interval that is much shorter than p and q .

Above, we found a sublattice of \mathcal{L}' with short basis vectors. One may wonder if there are short vectors in \mathcal{L}' that are not in the sublattice generated by these short vectors. To answer this question, we apply the Gaussian heuristic to the lattice that is obtained when the sublattice is projected out as in Section 6.1 of [10]. Lemma 5 on page 29 of [6] gives the explicit formula for the volume of a lattice resulting as the orthogonal projection over a linear subspace. We write it here for the convenience of the reader.

Lemma 2. *Let L be a d -dimensional lattice in \mathbb{R}^s and M be a r -dimensional sublattice of L which the property that one of its basis can be extended to a basis of L . Let π_M denote the orthogonal projection over the orthogonal supplement of the linear span of M . Then the image of L by π_M is a $(d-r)$ -dimensional lattice of \mathbb{R}^s and volume $\text{Vol}(L)/\text{Vol}(M)$.*

Assuming $p < q$ and $K > \alpha$, the volume of the resulting lattice equals

$$q^{c-\alpha-1} / \sqrt{\det(BB^t)},$$

where B is the matrix

$$B = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_c \\ \binom{x_1}{2} & \binom{x_2}{2} & \cdots & \binom{x_c}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \binom{x_1}{\alpha} & \binom{x_2}{\alpha} & \cdots & \binom{x_c}{\alpha} \end{pmatrix}.$$

When x_1, \dots, x_c are uniformly drawn from $[0, w)$, $\sqrt{\det(BB^t)}$ will be of order $w^{\alpha(\alpha+1)/2}$. Comparing powers of w , the resulting volume is therefore expected to be much larger than 1 if $K(c - \alpha) > \alpha(\alpha + 1)/2$.

Example. Let $\alpha = 6$, $w = 2^{16}$, and

$$p = 322503631145131659181549502994177879533$$

$$q = 322503631145131659181549502996361408083$$

so that $p \approx q \approx w^8$, and $K = 8$.

Suppose the coefficients of $f \in \mathbb{Z}_p[x]$ and $g \in \mathbb{Z}_q[x]$ are equal to

$$f_0 = 192299855391930388766069561100536978455$$

$$f_1 = 80324299466086676640269450973128212279$$

$$f_2 = 134802655995538131612821059755185358806$$

$$f_3 = 223036273860653058471857675170774765711$$

$$f_4 = 81615146624468266057406642183853219751$$

$$f_5 = 282812473825451509017913772106035640705$$

$$f_6 = 278906905917307720980382059680001096297$$

and

$$g_0 = 81564018199971421800339434244552477506$$

$$g_1 = 12324696623153181384549093381069011068$$

$$g_2 = 80030936209387920933656861269029654371$$

$$g_3 = 315635911037272927490950126509525457405$$

$$g_4 = 217950416300798270685940703747161570332$$

$$g_5 = 75454198535432609870859677101539890163$$

$$g_6 = 26892964982895845277700750286746366172.$$

In this example the smallest value of c for which $K(c - \alpha - 1) > \alpha(\alpha + 1)/2$ is 10. That means that if we pick $c = 10$ points uniformly from $[0, w)$, there is a fair chance that the volume of the projected lattice is much larger than 1.

We are given the values of $h(x) = \langle f(x) \rangle_p + \langle g(x) \rangle_q$ in $c = 10$ points randomly chosen from the interval $[0, w)$ according to the following table.

i	x_i	$h(x_i)$
1	34915	357083778061836956769804023406098677550
2	30844	501434122478371565756095361502998185705
3	55453	362669734592545590446623074678041228580
4	43386	453528102619044436291771088280150310990
5	61725	409617140945520234057946967178875528708
6	39144	426802401636630448727954157743588116409
7	14608	311556461063783252602939114845129657070
8	24287	594980681560119885662989234834546277705
9	24582	119430230752341918846040173886171897211
10	36432	20159634491993343981036574887019110187

Constructing the lattice as described before, including an additional row of ones in the matrix \mathbf{V} in order to take the constant terms of the polynomials into account, we find a lattice vector that is close to the target vector. The polynomial coefficients corresponding to this nearby lattice vector are

$$\begin{aligned}
\tilde{f}_0 &= 136931826884319377850275846232659046764 \\
\tilde{f}_1 &= 127274522470810992144873423947517028220 \\
\tilde{f}_2 &= 166540029496138250784732903087691991725 \\
\tilde{f}_3 &= 149982375974823828230059543913714128152 \\
\tilde{f}_4 &= 157228597180650695773338976918720767558 \\
\tilde{f}_5 &= 159036774649843108350794952315211705687 \\
\tilde{f}_6 &= 151078581747150708184414679388065540292
\end{aligned}$$

and

$$\begin{aligned}
\tilde{g}_0 &= 136932046707582432716133148636421185297 \\
\tilde{g}_1 &= 126626289190994695470719871904637347436 \\
\tilde{g}_2 &= 155794773090498354822261518935095270543 \\
\tilde{g}_3 &= 169208171060443111900860401561223641003 \\
\tilde{g}_4 &= 146816182843853780680863223220064874839 \\
\tilde{g}_5 &= 149958509619423673718575100602091084672 \\
\tilde{g}_6 &= 150242072053814918362813276371264593533.
\end{aligned}$$

The difference $\tilde{h}(x) - h(x)$ is plotted in Figure 1. This shows that, even though the interpolation is not perfect, it still predicts the correct value in a sizable fraction of the points and the error pattern does not look random.

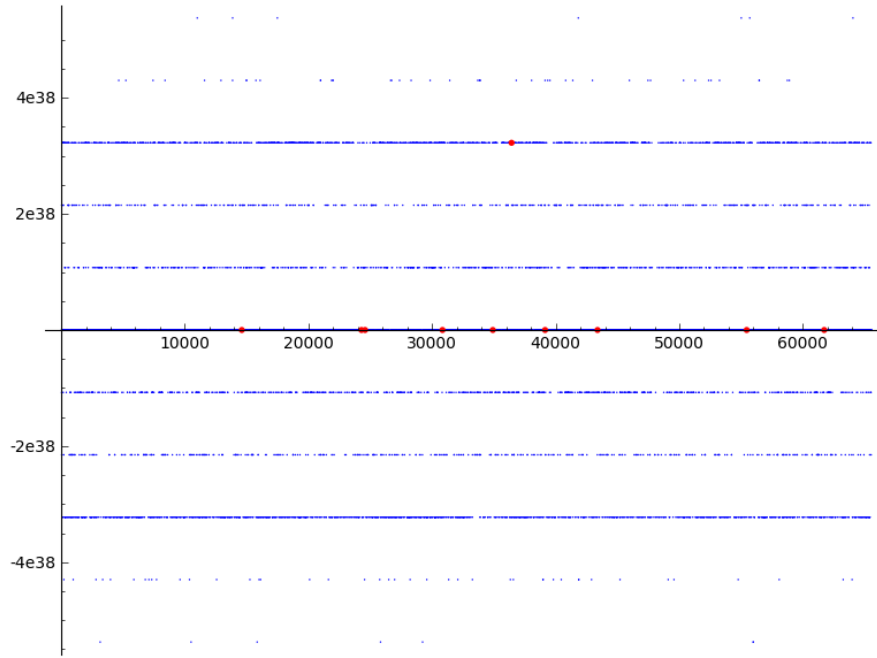


Figure 1: Graph of $\tilde{h}(x) - h(x)$. The reconstructed function $\tilde{h}(x)$ fits the observation perfectly in 9 out of the 10 points, but more interestingly, the interpolation error is zero in many other points, even though $c < 2\alpha$. If the error is non-zero, it is restricted to very narrow bands.

5. Conclusions

We have introduced the MMO problem. It seems infeasible to solve the MMO problem for unknown moduli. We have shown the equivalence of the MMO problem to finding close vectors in a lattice. If all observed function arguments lie in an interval that is much shorter than the moduli, then reconstruction of the unknown polynomials is infeasible; however, the computed polynomials often gives correct interpolation of the function on that short interval.

The MMO problem can readily be generalized to more than two moduli. Furthermore, an additional modular operation may be performed on the sum of the polynomial evaluations, see [2].

References

- [1] J. Boyar, “Inferring sequences produced by pseudo-random number generators,” *Journal of Association of Computing Machinery*, vol. 36, no. 1, pp. 129–141, 1989.
- [2] O. García-Morchoń, L. Tolhuizen, D. Gómez, and J. Gutierrez, “Towards fully collusion-resistant id-based establishment of pairwise keys,” *IACR Cryptology ePrint Archive*, vol. 2012, p. 618, 2012.
- [3] D. Micciancio and S. Goldwasser, *Complexity of lattice problems*, ser. The Kluwer International Series in Engineering and Computer Science, 671. Boston, MA: Kluwer Academic Publishers, 2002, a cryptographic perspective.
- [4] P. M. Gruber and C. G. Lekkerkerker, Eds., *Geometry of numbers*, ser. The Kluwer International Series in Engineering and Computer Science. Amsterdam: North-Holland Publishing Co., 1987, vol. 37. [Online]. Available: <http://dx.doi.org/10.1007/978-3-642-02295-1>
- [5] M. Grötschel, L. Lovász, and A. Schrijver, *Geometric algorithms and combinatorial optimization*. Berlin: Springer-Verlag, 1993.
- [6] P. Q. Nguyen and B. Vallée, Eds., *The LLL algorithm*, ser. Information Security and Cryptography. Berlin: Springer-Verlag, 2010, survey and applications. [Online]. Available: <http://dx.doi.org/10.1007/978-3-642-02295-1>

- [7] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, "Factoring polynomials with rational coefficients," *Math. Ann.*, vol. 261, no. 4, pp. 515–534, 1982. [Online]. Available: <http://dx.doi.org/10.1007/BF01457454>
- [8] I. Shparlinski and A. Winterhof, "Noisy interpolation of sparse polynomials in finite fields," *Appl. Algebra Eng. Commun. Comput.*, vol. 16, no. 5, pp. 307–317, Nov. 2005. [Online]. Available: <http://dx.doi.org/10.1007/s00200-005-0180-1>
- [9] L. Babai, "On lovász' lattice reduction and the nearest lattice point problem," *Combinatorica*, vol. 6, no. 1, pp. 1–13, 1986.
- [10] O. García-Morchón, R. Rietman, I. E. Shparlinski, and L. Tolhuizen, "Interpolation and approximation of polynomials in finite fields over a short interval from noisy values," *arXiv.org preprint archive*, vol. 2014, 2014. [Online]. Available: <http://arxiv.org/abs/1401.1331>